Draft genome sequence data of Lactobacillus paracasei strain DTA83 isolated from infant stools

Lemos Junior, Wilson José Fernandes; Fioravante Guerra, André; da Silva Duarte, Vinícius; Treu, Laura; Tarrah, Armin; Campanaro, Stefano; Luchese, Rosa Helena; Giacomini, Alessio; Corich, Viviana

Published in:
Data in Brief

Link to article, DOI:
10.1016/j.dib.2019.01.041

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Data in Brief 22 (2019) 1064–1067

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Draft genome sequence data of *Lactobacillus paracasei* strain DTA83 isolated from infant stools

Wilson José Fernandes Lemos Junior [a,1], André Fioravante Guerra [b,1], Vinícius da Silva Duarte [c], Laura Treu [d,e], Armin Tarrah [c], Stefano Campanaro [e], Rosa Helena Luchese [f], Alessio Giacomini [c,g,*], Viviana Corich [c,g]

[a] *Department of Biotechnology, University of Verona, Verona, Italy*
[b] *Federal Center of Technological Education Celso Suckow da Fonseca, Valença, RJ, Brazil*
[c] *Department of Agronomy Food Natural Resources Animals and Environment, University of Padova, Legnaro, Italy*
[d] *Department of Environmental Engineering, Technical University of Denmark, Kongens Lyngby, Denmark*
[e] *Department of Biology, University of Padova, Padova, Italy*
[f] *Department of Food Technology, Institute of Technology, Federal Rural University of Rio de Janeiro, RJ, Brazil*
[g] *Interdepartmental Centre for Research in Viticulture and Enology, University of Padova, Conegliano, Italy*

ARTICLE INFO

Article history:
Received 26 September 2018
Received in revised form
8 January 2019
Accepted 18 January 2019
Available online 19 January 2019

ABSTRACT

Here the draft genome sequence of *Lactobacillus paracasei* strain DTA83, isolated from stools of healthy infants in Rio de Janeiro (Brazil), is reported. The 2.8-Mb genome possesses 2825 protein-coding sequences distributed on 330 SEED subsystems. This strain belongs to a set of potentially probiotic *Lactobacillus* spp. strains used to study genetic factors related to antibiotic resistance after stress conditions, such as simulated gastrointestinal conditions. The complete genome data have been deposited in GenBank under the accession number QRBH00000000, https://www.ncbi.nlm.nih.gov/nuccore/ QRBH00000000.

* Corresponding author at: Department of Agronomy Food Natural Resources Animals and Environment, University of Padova, Legnaro, Italy.
E-mail address: alessio.giacomini@unipd.it (A. Giacomini).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.dib.2019.01.041
2352-3409/

Specification table

Subject area	Biology
More specific subject area	Microbiology, Genomics
Type of data	Genomic sequence, gene prediction of *Lactobacillus paracasei*
How data was acquired	Shotgun whole-genome DNA sequencing using MiSeq platform (Illumina) at the Ramaciotti Centre, Sydney, Australia
Data format	Raw draft genome assembly and gene prediction
Experimental factors	DNA was extracted from *Lactobacillus paracasei* DTA83
Experimental features	Isolation of *Lb. paracasei* DTA83, whole genome shotgun sequencing followed by assembly and annotation
Data source location	The strain *Lb. paracasei* DTA83 was isolated from stools of infants between 7 and 21 days old at Fernandes Figueira Institute of FIOCRUZ, Rio de Janeiro, Brazil.
Data accessibility	The Whole Genome Shotgun project of *Lb. paracasei* DTA83 has been deposited in DDBJ/ENA/GenBank under the accession no. QRBH00000000. https://www.ncbi.nlm.nih.gov/nuccore/ QRBH00000000
Related research article	A.F. Guerra, W.J.F. Lemos Junior, G.O. dos Santos, C. Andrighetto, A. Gianomini, V. Corich, R.H. Luchese, A.F. Guerra, W.J.F. Lemos Junior, G. O. dos Santos, C. Andrighetto, A. Gianomini, V. Corich, R.H. Luchese, *Lactobacillus paracasei* probiotic properties and survivability under stress-induced by processing and storage of ice cream bar or ice-lolly, Ciência Rural. 48 (2018). doi:10.1590/0103-8478cr20170601. [10]

Value of the data

- The compete genome of *Lb. paracasei* DTA83, isolated from infant stools, allows to study in depth the genetic information of this strain in view of its potential use as probiotic strain.
- The sequence data will be useful for comparative studies towards environmental strains or others of intestinal origin.
- The availability of genome sequence can allow the discovery of other technological or health related potentially relevant characteristics of this strain.

1. Data

Lactobacillus paracasei is a lactic acid bacterium (LAB) with interesting technological potentialities, mostly related to its capability to acidify the substrate, since it is known that a pH decrease affects bacterial [1,2] and also yeast [3,4] population composition and dynamics.

Assembly of the shotgun reads generated 70 scaffolds, which were ordered and oriented with CONTIguator [5], using strain *Lb. paracasei* ATCC 334 as reference, giving a total genome size of 2.8 Mb, with a GC content of 46.4%. The average nucleotide identity (ANI) of 98.56% with *Lb. paracasei* ATCC 334, confirms that strain DTA83 belongs to the *Lb. paracasei* species [6]. The Rapid Annotations using Subsystems Technology (RAST) server [7] was used for gene prediction and annotation. The number of predicted protein-coding sequences (CDSs) in DTA83 was 2825, distributed on 330 SEED subsystems. Moreover, 59 structural RNAs were found. In total, 31 genes are related to the “resistance to antibiotics and toxic compounds” subsystem. Among these, it must be highlighted the presence of four genes associated to fluoroquinolones resistance, a feature not identified in the reference strain *Lb. paracasei* ATCC 334. PHASTER [8] analysis returned three prophage regions (1 intact, 1 incomplete

and 1 questionable). Additionally, one confirmed two “questionable” clusters of regularly interspaced short palindromic repeats (CRISPRs) are present in this strain, as predicted by CRISPRFinder [9].

Data is available at https://www.ncbi.nlm.nih.gov/nuccore/ QRBH00000000.

2. Experimental design, materials and methods

The strain used in this work was isolated from the faeces of healthy infants aged between 7 to 21 days, assisted by the Human Milk Bank (HMB) and the Neonatal Intensive Care Unit (NICU) of the Neonatal Fernandes Figueira Institute of Oswaldo Cruz Foundation (FIOCRUZ), located in Rio de Janeiro, Brazil [10]. Approximately 1 g of stools was collected, serially diluted in Wilkins-Chalgren broth (Oxoid, Basingstoke, United Kingdom) and plated on LAMVAB agar [11]. Plates were incubated at 37 °C for 48 h under partial anaerobic conditions and characteristic colonies were selected and subjected to biochemical and genotypic tests.

For DNA extraction, DTA83 was grown for 24 h at 37 °C on de Man Rogosa Sharp (MRS) broth. Afterwards, cells were harvested (5000 rpm for 5 min at 4 °C) and high-quality genomic DNA was extracted using the PowerFood™ Microbial DNA isolation kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer instructions. DNA quantification was performed using both NanoDrop (ThermoFisher Scientific, Waltham, MA, USA) and Qubit (ThermoFisher Scientific).

Genome sequencing was carried out with an Illumina MiSeq sequencer and Nextera XT libraries at the Ramaciotti Centre (Sydney, Australia). The average number of paired-end reads (2×250 bp) for the strain was 3,034,380.

The completed genome data was deposited in NCBI GenBank under the accession number QRBH00000000, https://www.ncbi.nlm.nih.gov/nuccore/ QRBH00000000.

Acknowledgements

This work has been funded in part by the Ministero dell’Università e della Ricerca Scientifica (MIUR), Italy. WJFLJ was financially supported by CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior., Brasil.

Transparency document. Supplementary material

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.01.041.

References

[1] P.A. Maragkoudakis, T. Nardi, B. Bovo, M. D'Andrea, K.S. Howell, A. Giacomini, V. Corich, Biodiversity, dynamics and ecology of bacterial community during grape marc storage for the production of grappa, Int. J. Food Microbiol. 162 (2013) 143–151. https://doi.org/10.1016/j.ijfoodmicro.2013.01.005.

[2] G. Giraffa, Studying the dynamics of microbial populations during food fermentation: table 1, FEMS Microbiol. Rev. 28 (2004) 251–260. https://doi.org/10.1016/j.femsre.2003.10.005.

[3] B. Bovo, A. Giacomini, V. Corich, Effects of grape marcs acidification treatment on the evolution of indigenous yeast populations during the production of grappa, J. Appl. Microbiol. 111 (2011) 382–388. https://doi.org/10.1111/j.1365-2672.2011.05060.x.

[4] B. Bovo, T. Nardi, F. Fontana, M. Carlot, A. Giacomini, V. Corich, Acidification of grape marc for alcoholic beverage production: effects on indigenous microflora and aroma profile after distillation, Int. J. Food Microbiol. 152 (2012) 100–106. https://doi.org/10.1016/j.ijfoodmicro.2011.10.006.

[5] M. Galardini, E.G. Biondi, M. Bazzicalupo, A. Mengoni, CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes, Source Code Biol. Med. 6 (2011) 11. https://doi.org/10.1186/1751-0473-6-11.

[6] L.M. Rodriguez-R, K.T. Konstantinidis, Bypassing cultivation to identify bacterial species, Microbe Mag. 9 (2014) 111–118. https://doi.org/10.1128/microbe.9.111.1.
[7] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST server: rapid annotations using subsystems technology, BMC Genom. 9 (2008) 75. https://doi.org/10.1186/1471-2164-9-75.
[8] D. Arndt, J.R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, D.S. Wishart, PHASTER: a better, faster version of the PHAST phage search tool, Nucleic Acids Res. 44 (2016) 1–6. https://doi.org/10.1093/nar/gkw387.
[9] I. Grissa, G. Vergnaud, C. Pourcel, The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats, BMC Bioinform. 8 (2007) 172. https://doi.org/10.1186/1471-2105-8-172.
[10] A.F. Guerra, W.J.F. Lemos Junior, G.O. dos Santos, C. Andrighetto, A. Gianomini, V. Corich, R.H. Luchese, A.F. Guerra, W.J. F. Lemos Junior, G.O. dos Santos, C. Andrighetto, A. Gianomini, V. Corich, R.H. Luchese, *Lactobacillus paracasei* probiotic properties and survivability under stress-induced by processing and storage of ice cream bar or ice-lolly, Ciênc. Rural (2018), https://doi.org/10.1590/0103-8478cr20170601.
[11] R. Hartemink, V. Domenech, F. Rombouts, LAMVAB—a new selective medium for the isolation of lactobacilli from faeces, J. Microbiol. Methods (1997), https://doi.org/10.1016/S0167-7012(97)00025-0.